FluxSat README FILE

Overview

This document presents a brief description of the FluxSat data products. These are products produced by PI Joanna Joiner. FluxSat (version 2.0) is derived from the MODerate-resolution Imaging Spectroradiometer (MODIS) instruments on the NASA Terra and Aqua satellites using the MCD43C Bidirectional Reflectance Distribution Function (BRDF)-Adjusted Reflectances (NBAR) (Schaaf, 2015). Output is "calibrated" using a set of the FLUXNET 2015 eddy covariance data and has been compared with independent data (i.e., not used in the calibration) as validation. Global Gross Primary Production (GPP) estimates are currently available from March 2000 through present on a best effort basis.

Data Quality Assessment

Users should be aware that the data sets provided here have undergone only a limited amount of validation. FluxSat v1.0 GPP estimates were also compared with other satellite driven data sets such as FLUXCOM and the Vegetation Photosynthesis Model (VPM). Please see Joiner *et al.* (2018) for more details. FluxSat GPP v1.0 data showed less bias and higher precision than other data sets with respect to independent FLUXNET 2015. FluxSat v2.0 was shown to outperform v1.0 (Joiner and Yoshida, 2020).

Known Algorithm Features:

1) The product is only as good as the input MODIS MCD43C reflectances. Any issues with those reflectances will alias into the final GPP product.

2) We are providing data for the pre-Aqua period in which only the Terra satellite is used. We have found that the data

quality, particularly in the partial year 2000 is not as good as the period from 2003 (Aqua and Terra) onwards. Users are cautioned when using the Terra-only period data.

3) We find that there is a very small fraction of spurious values that can be seen as discontinuities in time series. These data points are usually associated with low percent of data as seen in the "Percent_Inputs" variable (typically < 30%) and are associated with a BRDF_Quality value of 5. However, because there are also many good data points with BRDF quality 5, we have left all values in the data set. Users are cautioned when using data with low percent inputs.

4) A gridbox GPP uncertainty is derived for each point based on machine learning. The uncertainty is conservative in that it includes representativeness error in the training set (see Joiner and Yoshida, 2020), but note that errors may be spatially and/or temporally correlated.

Product Description

The FluxSat gridded products are written as a self-describing NetCDF files. The FluxSat v2.0 native GPP product is computed at the spatio-temporal resolution of the MCD43C data set (daily at 0.05° spatial resolution, i.e., the so-called Climate Modeling Grid, CMG). Note that MCD43C is produced with a rolling 16 day, so data are not truly daily, but are weighted towards the day of interest. A simple IDL reader is provided. The information provided on these files includes: Latitude, longitude, GPP, GPP uncertainty, an estimate of the FPAR*$LUE_{constitutive}$ (scaled to produce values in approximately the zero to one range, see Joiner and Yoshida for further explanation), and an expanded quality flag and percent inputs from the MCD43C data set.

Changes from version 1

1. Version 2.0 uses machine learning with MODIS reflectances trained on FLUXNET 2015 data rather than a linear combination of bands (similarly trained) used in v1.0 (see Joiner and Yoshida, 2020). Note that a different training is used in FluxSat v2.0 as compared with what is shown in Joiner and Yoshida (2020). Here, the training was conducted with MCD43C rather than the 1km MCD43D reflectances. All sites used in Joiner and Yoshida (2020) are used in the final training. Meteorological data are not used in FluxSat v2.0, only MODIS bands 1-7 are used along with an estimate of top-of-atmosphere PAR.

2. GOME-2 SIF was used in v1.0. We found that the low spatial resolution of this data set produced some blocky spatial patterns in the high-resolution data. With machine learning, we found that it was no longer necessary to use GOME-2 SIF, so it is not used in v2.0.

3. Version 2.0 is gap-filled using a climatology developed from the full data set. Gap-filled data are noted within the BRDF quality flag. GPP uncertainties for gap-filled data are set to 2 g C m^{-2} day^{-1}. When daily PAR is very low (typical of dark winter days), GPP gap-filled values and uncertainties are set to zero.

4. Quality assurance data (BRDF_Quality and Percent_Inputs) are now carried over from the MCD43C data set.

Contact

All questions related to the FluxSat datasets should be directed to Joanna Joiner (Joanna.Joiner@nasa.gov). It would be helpful to send the PI a copy of any publication that uses these data for tracking purposes.

Acknowledgments

We gratefully acknowledge the MODIS data processing teams, particularly the producers of MCD43 (C. Schaaf et al.) for providing the data used here.

References

Joiner, J., **Y. Yoshida**, Y. Zhang, et al. 2018. "Estimation of Terrestrial Global Gross Primary Production (GPP) with Satellite Data-Driven Models and Eddy Covariance Flux Data." *Remote Sensing*, **10 (9):** 1346 [10.3390/rs10091346]

Joiner, J., and **Y. Yoshida**. 2020. "Satellite-based reflectances capture large fraction of variability in global gross primary production (GPP) at weekly time scales." *Agricultural and Forest Meteorology*, **291:** 108092 [10.1016/j.agrformet.2020.108092]

Schaaf, C., 2015. MCD43D62 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF-Adjusted Band1 daily L3 global 30ArcSec CMG V006. NASA EOSDIS Land Processes DAAC. 10.5067/modis/mcd43D62.006.